

文章编号 1004-924X(2023)15-2295-24

## 基于深度学习的遥感图像舰船目标检测算法综述

黄泽贤<sup>1,2</sup>, 吴凡路<sup>1</sup>, 傅瑶<sup>1</sup>, 张雨<sup>1</sup>, 姜肖楠<sup>1\*</sup>

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国科学院大学, 北京 100049)

**摘要:**海面舰船目标检测是遥感图像处理 and 模式识别领域备受关注的重点研究方向,对舰船目标的自动检测在民用和军用方面都具有重大意义。梳理和分析了典型基于深度学习的目标检测算法的优缺点,并进行了对比和总结;归纳了基于深度学习的舰船目标检测的技术现状,并从多尺度检测、多角度检测、小目标检测、模型轻量化和大幅宽遥感图像舰船目标检测等方面对技术现状进行了详细的介绍。最后,介绍了舰船目标识别算法常用的评价标准和现有的舰船图像数据集,探讨了遥感图像舰船目标检测算法现在所面临的问题和未来的发展趋势。

**关键词:** 遥感图像;舰船目标检测;卷积神经网络;图像数据集

中图分类号: TP391.4 文献标识码: A doi: 10.37188/OPE.20233115.2295

## Review of deep learning-based algorithms for ship target detection from remote sensing images

HUANG Zexian<sup>1,2</sup>, WU Fanlu<sup>1</sup>, FU Yao<sup>1</sup>, ZHANG Yu<sup>1</sup>, JIANG Xiaonan<sup>1\*</sup>

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

\* Corresponding author, E-mail: jxn\_ciomp@qq.com

**Abstract:** The detection of naval targets is a key area of research interest in the field of remote sensing image processing and pattern recognition. Moreover, the automatic detection of naval targets is crucial to both civil and military applications. In this study, we discuss and analyze the advantages and disadvantages of typical deep-learning-based target-detection algorithms, compare and summarize them, and summarize state-of-the-art deep-learning-based ship target detection methods. We also provide a detailed introduction to five aspects of state-of-the-art ship target detection methods, including multi-scale detection, multi-angle detection, small target detection, model light-weighting, and large-format wide remote sensing imaging. We also introduce the common evaluation criteria of ship target recognition algorithms and existing ship image datasets, and discuss the current problems faced by ship target detection algorithms using remote sensing images and future development trends in the field.

**Key words:** remote sensing imagery; ship target detection; convolutional neural networks; image dataset

收稿日期: 2022-09-16; 修订日期: 2022-12-09.

基金项目: 国家自然科学基金青年科学基金资助项目 (No. 42001345)

## 1 引言

我国拥有辽阔的海域,海面舰船目标检测无论是在民用方面还是军事方面都具有十分重要的意义。遥感图像目标检测识别一直是遥感图像处理 and 模式识别领域备受关注的研究方向<sup>[1-2]</sup>。基于遥感图像的舰船目标检测技术使得大范围远海域的监测成为可能,极大地丰富军事、海事部门的监测手段。随着遥感卫星技术的不断成熟,遥感图像分辨率不断提升、数据规模日益猛增,传统目标检测算法主要基于手工提取特征,手工提取特征存在识别准确率不高、效率低、易受背景干扰等缺点<sup>[3-4]</sup>,已难以满足应用需求。

在计算机视觉领域,基于深度学习的方法相较于传统方法显示出巨大优势。深度学习方法可以从海量图像数据中学习图像特征表达以极大地提高含有大量信息的图像处理精度。通过组合多个非线性变换、自适应地组合低层特征形成更抽象的高层特征的深度网络,进而提取出图像中的光谱、纹理、几何等隐藏得更深、语义信息

更丰富的特征,获得比传统方法更高的精度和效率<sup>[5]</sup>。本文归纳梳理了经典目标检测算法,对遥感图像舰船目标检测算法的技术现状进行了分析,探讨了当前遥感图像船舶目标检测算法面临的问题与挑战以及未来的发展趋势。

## 2 深度学习目标检测算法

随着深度学习在计算机视觉领域取得的革命性成功,卷积神经网络(Convolutional Neural Networks, CNN)被广泛应用于图像分类与目标识别。CNN可以自动提取特征,大大提高了目标识别的准确率。基于深度学习的目标检测算法主要分为基于锚框的(Anchor-Based)方法和无锚框的(Anchor-Free)方法,其中Anchor-Based方法包括基于候选区域(Region Proposals)的两阶段目标检测算法和基于回归的一阶段目标检测算法。基于深度学习的目标检测算法的发展进程见图1,本文介绍了几种经典的目标检测算法。

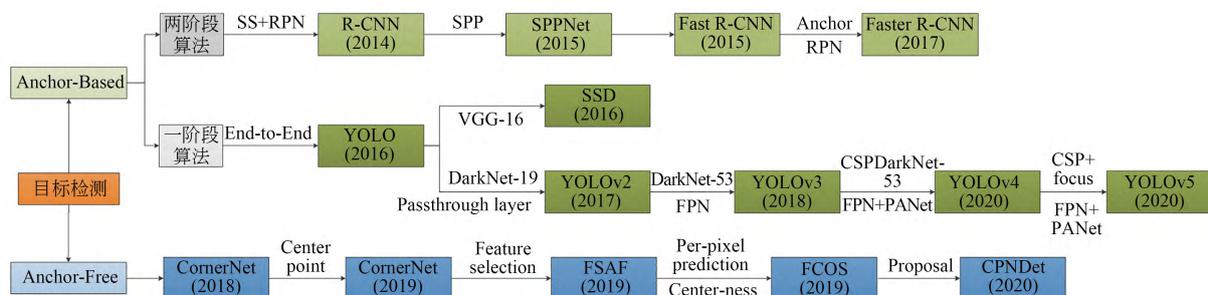


图1 目标检测算法发展进程

Fig. 1 Development process of target detection algorithms

### 2.1 Anchor-based 目标检测算法

#### 2.1.1 基于候选区域的两阶段目标检测算法

两阶段检测算法首先从图像中提取候选区域,然后从候选区域生成目标的预测框。两阶段检测算法一般检测精度较高,但检测速度慢。

##### (1) R-CNN

Girshick提出的区域卷积神经网络(Region Convolutional Neural Networks, R-CNN)<sup>[6]</sup>首先采用选择性搜索(Selective Search, SS)算法获取可能包含目标的建议区域;然后,将建议区域的尺寸调整一致后送入CNN AlexNet中提取特征;

再将提取的特征向量送入每类的支持向量机(Support Vector Machine, SVM)分类器进行二分类,判断目标是否属于该类;最后对已分类的目标框进行精细调整得到更加准确的边界框坐标。检测流程见图2。

R-CNN算法第一次将卷积神经网络用于目标检测,在VOC2012(Visual Object Classes Challenge 2012)数据集<sup>[7]</sup>上的均值平均精度(mean Average Precision, mAP)达到了53.3%,相比之前的最优结果提高了30%以上。但该算法一张图像生成大量的候选框,特征的冗余计算

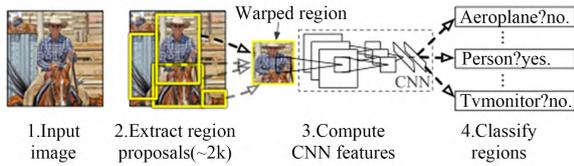


图 2 R-CNN 检测流程

Fig. 2 R-CNN detection process

使得检测速度很慢。

(2)SPPNet

空间金字塔池化层 (Spatial Pyramid Pooling, SPP)<sup>[8]</sup>将一幅图像分为若干个尺度的图像块,对提取的特征进行融合得到多尺度特征。SPPNet能接收不同尺寸的输入并生成尺寸一致的特征图,网络结构如图 3 所示。

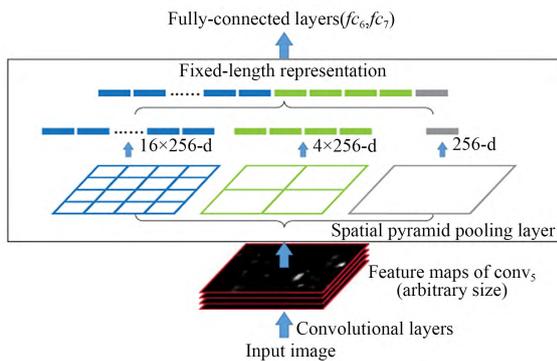


图 3 SPPNet 网络结构

Fig. 3 SPPNet network structure

SPPNet解决了CNN模型输入图像尺寸固定的问题,避免特征图的重复计算,在VOC2007数据集<sup>[9]</sup>上的mAP值为59.2%,比R-CNN的检测速度快。但是,和R-CNN相同,SPPNet训练CNN提取特征后送入SVM进行分类,耗费巨大的存储空间,多阶段训练流程复杂,而且不能微调空间金字塔池化之前的卷积层,限制了深层网络的准确性。

(3)Fast RCNN

Fast R-CNN<sup>[10]</sup>可以同时训练分类器和边框回归器。Fast R-CNN首先使用SS算法生成候选区域;然后,将图像输入至VGG-16 (Visual Geometry Group Network)<sup>[11]</sup>提取特征,得到感兴趣区域 (Region of Interest, ROI);再在ROI上利用池化层将特征图缩放到相同尺寸,最后将这些特征图传递到全连接层进行分类,并用Soft-

max和线性回归层得到目标边界框,架构如图4所示。

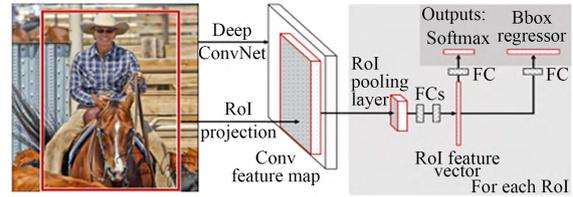


图 4 Fast R-CNN 架构

Fig. 4 Fast R-CNN structure

Fast R-CNN训练VGG-16网络的速度比R-CNN快9倍,测试速度快213倍,在VOC2012数据集上实现了更高的mAP (65.7%)。与SPPNet相比,Fast R-CNN对VGG-16的训练速度快3倍,测试速度快10倍,准确率也更高。但是Fast R-CNN仍然使用SS算法获取感兴趣区域,速度上无法满足实时检测的应用需求。

(4)Faster RCNN

Faster R-CNN<sup>[12]</sup>用区域选择网络 (Region Proposal Networks, RPN)取代SS算法生成候选框,提高了检测速度。Faster R-CNN首先将图像输入至VGG-16得到特征图;然后用RPN生成目标建议区域;再应用ROI池化层将特征图和目标推荐区域调整到相同尺寸,最后输入全连接层生成目标的预测边界框,架构如图5所示。

Faster R-CNN在VOC2012数据集上的mAP值为67.0%,精度更高,并且检测速度更

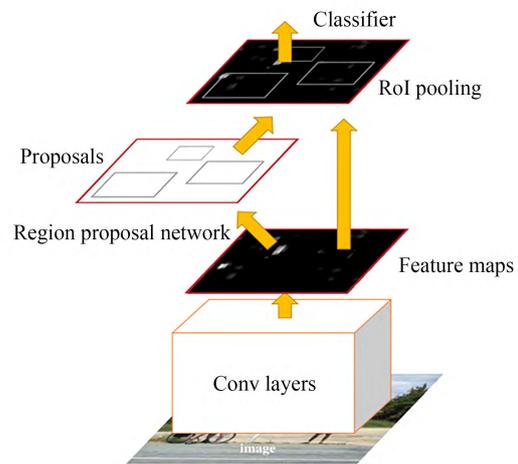


图 5 Faster R-CNN 架构

Fig. 5 Faster R-CNN structure

快,接近于实时检测,但是在后续检测阶段存在计算冗余。Faster R-CNN的主要缺点是交并比(Intersection Over Union, IOU)阈值过高会导致模型过拟合,过低则会产生噪声引起的虚警问题。

### 2.1.2 基于回归的一阶段目标检测算法

一阶段检测算法不需要生成候选区域,直接预测出目标的类别概率和位置信息。相比于两阶段目标检测算法,检测速度得到了很大的提升。

#### (1)YOLO

YOLO (You Only Look Once)<sup>[13]</sup>首次把目标检测看作一个回归问题,利用整张图像作为网络的输入,仅经过一个CNN,就可以得到边界框的位置及其所属的类别。YOLO将输入图像平均划分为 $S \times S$ 个网格,如果一个物体的中心落在某一个网格中,那么该网格负责检测该物体。每个网格要预测 $B$ 个边界框,每个边界框预测5个值:中心点坐标为 $(x, y)$ ,长宽 $(w, h)$ 和物体是否属于某个类别的置信度。此外每个网格还要预测类别信息,记为 $C$ 个类。一张图像输入网络输出一个 $S \times S \times (5 \times B + C)$ 的张量,网络结构如

图6所示。

相比于两阶段检测算法,YOLO的检测速度有很大的提高,能够以每秒45帧的速度实时处理 $448 \times 448$ 的图像;但是检测精度较低,在VOC2012数据集上的mAP值仅为57.9%,而且对小目标检测效果差。

#### (2)SSD

SSD (Single Shot MultiBox Detector)<sup>[14]</sup>同时借鉴了YOLO的回归思想和Faster R-CNN的锚机制,以VGG-16作为主干特征提取网络,在VGG-16之后添加了几个卷积层,利用低层特征与高层特征实现多尺度检测,网络结构如图7所示。SSD在多个特征层产生锚框,进行非极大值抑制(Non-Maximum Suppression, NMS)后输出最终检测结果。

在59帧/秒的处理速度下针对 $300 \times 300$ 大小的输入图像,SSD在VOC2007数据集上将mAP值提高至74.3%,性能优于Faster R-CNN。尽管SSD不同的检测分支可以对多个尺度的目标进行检测,但是用于小目标检测识别的低层特征仅有一层,并未对特征进行融合,特征表达能力不够、细节信息不足,造成部分小目标漏检。

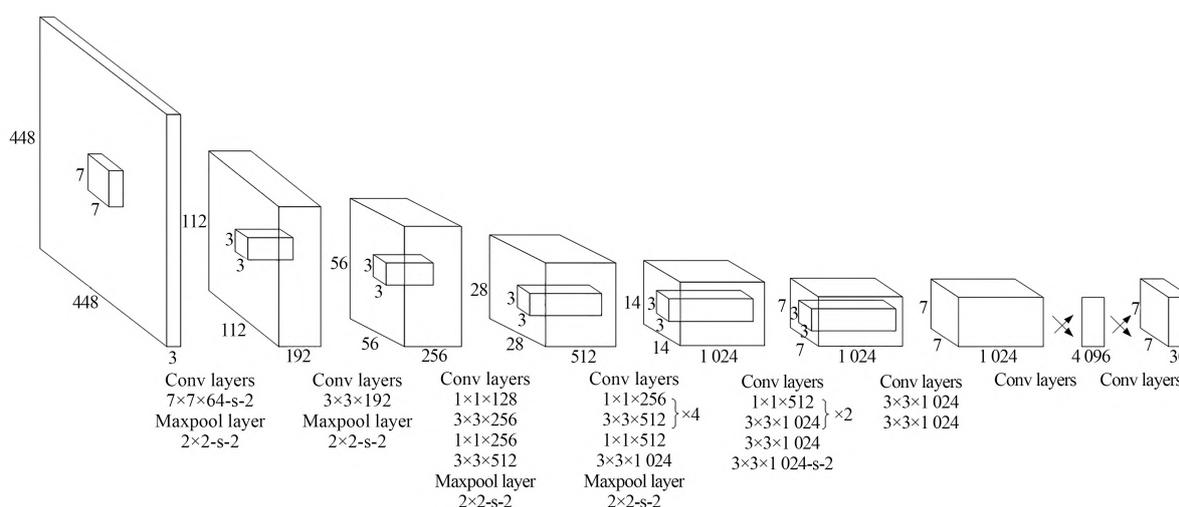


图6 YOLO网络结构

Fig. 6 YOLO structure

#### (3)YOLOv2

YOLOv2<sup>[15]</sup>针对YOLO召回率和定位精度方面的不足进行改进,检测速度更快。YOLOv2使用Darknet-19作为特征提取网络,可输入多种

尺寸的图像,每层卷积后增加批量标准化(Batch Normalization, BN)进行预处理;删除全连接层,引入了先验框来预测边界框坐标,并使用K-means聚类方法得到先验框的尺寸。还通过添加

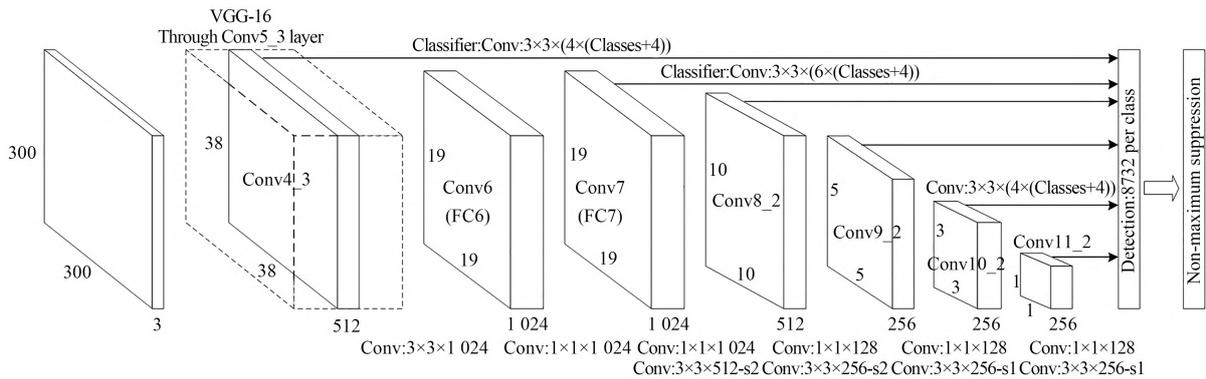


图7 SSD网络结构

Fig. 7 SSD structure

直通层,把高分辨率的浅层特征连接到低分辨率的深层特征而后进行融合获取细粒度特征,提高检测效果。输入  $416 \times 416$  大小图像时最终输出  $13 \times 13 \times N$  的特征图,  $N = (\text{class\_num} + 4 + 1) \times \text{anchor\_num}$ ; 其中  $\text{class\_num}$  为数据集中目标类别数目,  $\text{anchor\_num}$  是先验框数目。

在 67 帧/秒的处理速度下针对  $416 \times 416$  的输入图像, YOLOv2 在 VOC2007 数据集上将 mAP 值提高至 76.8%, 检测精度和速度均优于 SSD 和 Faster R-CNN。但是由于 YOLOv2 网络只有一条检测分支, 缺乏对多尺度上下文信息的获取, 对小目标的检测效果较差。

#### (4) YOLOv3

YOLOv3<sup>[16]</sup> 主干特征提取网络采用更深层的 Darknet-53, 利用特征金字塔网络结构 (Feature Pyramid Network, FPN) 进行特征融合实现了 3 个尺度的检测, 使用逻辑回归代替 softmax 进行多标签分类。YOLOv3 在兼顾实时性的同时保证了检测的准确性。

输入图像尺寸为  $320 \times 320$  时, YOLOv3 在 COCO 数据集<sup>[17]</sup> 上的平均精度 (Average Precision, AP) 为 28.2%, 单帧运行时间为 22 ms, 与 SSD 精度接近, 但速度快了三倍。但 YOLOv3 使用均方误差 (Mean Squared Error, MSE) 作为边界框回归损失函数, 使得目标的定位并不精准。

#### (5) YOLOv4

YOLOv4<sup>[18]</sup> 结合近年来 CNN 最优秀的优化策略对 YOLOv3 进行改进。YOLOv4 的主干特征提取网络为 CSPDarknet53, 使用 Mish 激活函数, 采用 SPP、路径聚合 PANet 作为加强特征提

取网络, 对特征进行融合来提升特征种类的多样性以及检测算法的鲁棒性。YOLOv4 还在数据预处理方面引入了 Mosaic 数据增强、cmBN (Cross mini-Batch Normalization) 和自对抗训练 (Self-Adversarial Training, SAT)。在预测阶段 YOLOv4 采用 CIU (Complete-IOU) 代替 MSE 作为边界框损失函数提高了定位精度, 同时将非极大值抑制 (Non Maximum Suppression, NMS) 换成 DIU\_NMS (Distance-IOU\_NMS), 避免相邻目标检测时出现漏检。

以 65 帧/秒的速度处理  $608 \times 608$  的输入图像, YOLOv4 在 COCO 数据集上的 AP 值为 43.5%, 实现了检测速度与精度的平衡。

#### (6) YOLOv5

YOLOv5<sup>[19]</sup> 在输入端利用 Mosaic 数据增强来提高小目标检测效果、训练前自动计算适合数据集的初始锚框, 并将图片缩放为统一尺寸。主干采用 Focus 结构和 CSP 结构, Focus 结构利用切片操作把输入的高分辨率特征图拆分为多个低分辨率特征图后再进行拼接后进行卷积得到输出特征图, Focus 可以减少参数量、提升检测速度。颈部采用 FPN 和 PAN 进行特征融合, 并使用借鉴 CSPNet 设计的 CSP2 结构来代替普通卷积来加强颈部的特征融合能力。预测阶段 YOLOv5 采用 GIOU 损失和 DIU\_NMS。

YOLOv5 共有 4 种网络结构: YOLOv5s, YOLOv5m, YOLOv5l 和 YOLOv5x, 随着网络深度的不断增加, 精度不断上升, 速度随之下降。以 50 帧/秒的速度处理  $640 \times 640$  的输入图像, YOLOv5x 在 COCO 数据集上的 AP 值为 55%。

### 2.1.3 小结

Anchor-Based 目标检测算法根据预先设定的锚框来调整预测结果,检测性能对于锚框的大小、数量和长宽比异常敏感。固定的锚框使得检测器的通用性很差,通过预先定义尺度、长宽比生成的锚框通常只适用于一个或几个特定的对象。当新数据集中的目标尺寸和形状或图像尺寸与原始数据集有较大差异时,需要重新设置尺度、长宽比以适应新的目标检测数据集。为了匹配目标的真实框,网络会生成大量的锚框,训练时大部分被标记为负样本,这样就会造成正负样本不均衡的问题,干扰算法的学习过程。此外,在训练过程中,网络会计算真实框和所有锚框的交并比(Intersection over Union, IOU)来确定用于检测真实目标的锚框,这会占用大量的内存,消耗大量的时间。

## 2.2 Anchor-Free 目标检测算法

Anchor-Based 目标检测算法由于生成的锚框过多导致检测过程复杂,同时产生的大量超参数也会影响检测器的性能,而 Anchor-Free 目标检测算法通过确定关键点代替锚框大大减少了超参数的数量。

### 2.2.1 CornerNet

CornerNet<sup>[20]</sup>首次提出 Anchor-Free 思想,把检测目标的边界框转化为检测左上角和右下角的一对关键点,无需设计锚框作为先验框,减少了网络的超参数,架构如图 8 所示。

在 COCO 数据集上 CornerNet 的 AP 值为 42.1%。由于 CornerNet 只关注边缘和角点,缺

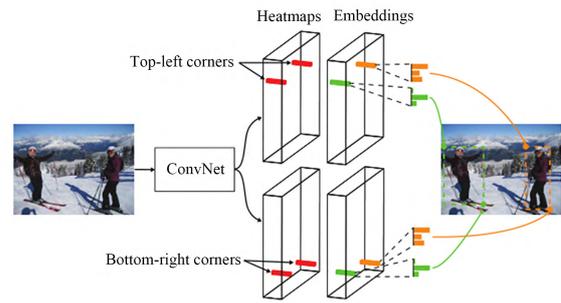


图 8 CornerNet 网络结构  
Fig. 8 CornerNet structure

乏目标内部信息,容易产生假正例(False Positives, FP),网络需要很多后处理(如 NMS)来得到预测结果,降低了算法检测速度。

### 2.2.2 CenterNet

Zhou 等在 CornerNet 基础上进行改进提出了 CenterNet 目标检测器<sup>[21]</sup>,架构如图 9 所示,直接检测目标中心和回归目标尺寸。该算法简单、快速、准确,不需要耗时的 NMS 后处理,具有端到端、可微分的特点,是真正意义上的 Anchor-Free。此外,CenterNet 具有良好的通用性,可以在一次前向传播中估计一系列额外的物体属性(如姿势、3D 方向、深度),可用于 3D 目标检测。

在 142 帧/秒的处理速度下 CenterNet 在 COCO 数据集上的 AP 值为 28.1%。但是 CenterNet 对于同一类别紧密相邻的目标检测效果较差,因为目标真实框的中心产生重叠,CenterNet 只能检测出一个中心点,造成目标漏检。

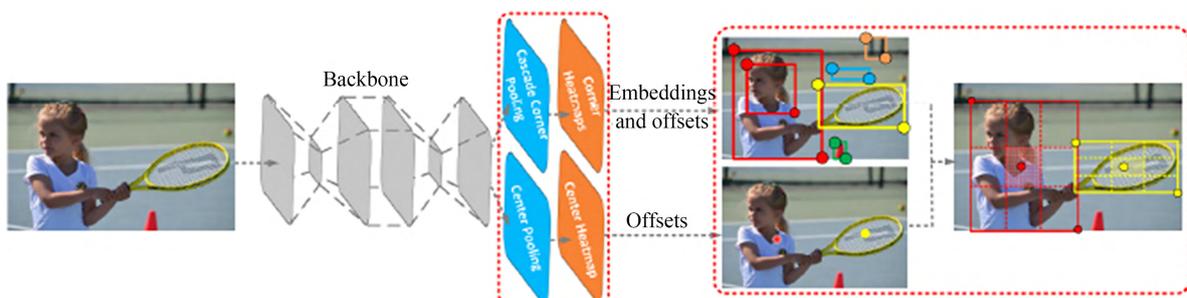


图 9 CenterNet 网络架构  
Fig. 9 CenterNet structure

### 2.2.3 FASF

Zhu 等提出的 FSAF (Feature Selective An-

chor-Free Module)模块<sup>[22]</sup>将在线特征选择用于训练特征金字塔中的无锚分支,为目标自动分配最

合适的特征,架构如图 10 所示。在推理时,FSAF 模块可以与基于锚的分支并行输出预测结果,几

乎不增加推理开销。引入 FSAF 模块的最佳模型在 COCO 数据集上的 AP 值为 44.6%。

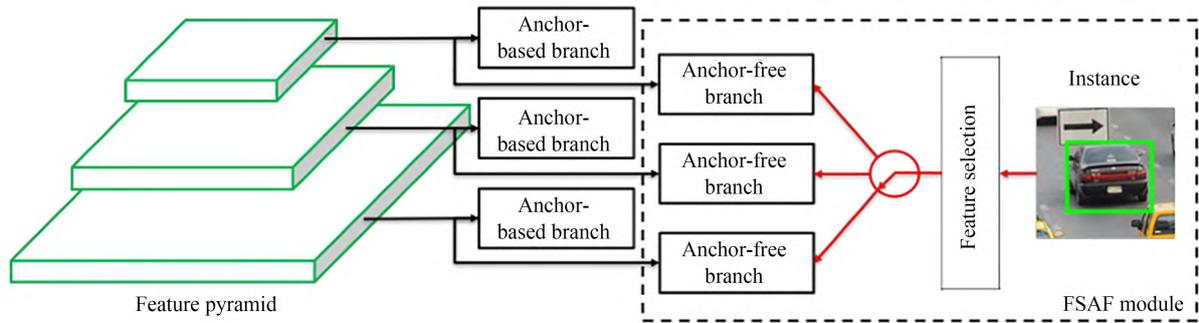


图 10 FSAF 模块

Fig. 10 FSAF modules

### 2.2.4 FCOS

Tian 等提出的 FCOS (Fully Convolutional One-Stage Object Detector)<sup>[23]</sup>以逐像素预测的方式进行目标检测,完全避免了与锚框相关的计算和超参数,网络结构如图 11 所示。通过引

入 FPN 用不同的层处理不同的目标框,解决目标真实框重叠时出现的漏检问题;同时引入了 Center-ness 层,过滤掉大部分的误检框。FCOS 检测器在 COCO 数据集上的 AP 值为 44.7%。

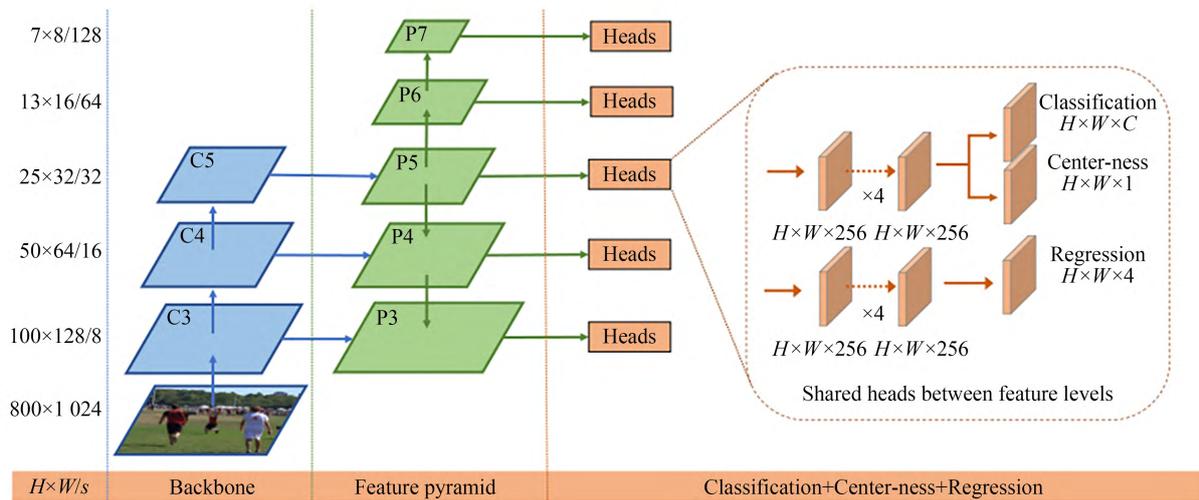


图 11 FCOS 架构

Fig. 11 FCOS architecture

上述基于关键点的方法消除了生成锚框的繁琐过程,通过直接根据网络预测关键点生成适合目标尺度和形状的方框,大大提高了检测速度。但是对关键点预测的准确性要求高,预测不准确容易导致漏检和定位不准。

### 2.3 算法性能比较

上述基于深度学习的目标检测算法在 VOC, COCO 等自然图像数据集上取得了较好的检测效果,各种算法的检测效果对比如表 1 所示。

表 1 经典算法检测效果对比

Tab.1 Comparison of classical algorithm detection results

数据集 算法	VOC2007 (mAP/%)	VOC2012 (mAP/%)	COCO (AP/%)
R-CNN	58.5	53.3	—
SPPNet	59.2	—	—
Fast R-CNN	66.9	65.7	19.7
Faster R-CNN	69.9	67.0	21.9
YOLO	63.4	57.9	—
SSD	74.3	72.4	28.8
YOLOv2	78.6	73.4	21.6
YOLOv3	—	—	28.2
YOLOv4	—	—	43.5
YOLOv5	—	—	55
CornerNet	—	—	42.1
CenterNet	—	—	28.1
FASF	—	—	44.6
FCOS	—	—	44.7

### 3 舰船目标检测技术现状

遥感图像目标检测作为目标检测的一个研究热点,一直受到研究者的广泛关注,经典目标检测算法如 Faster RCNN、YOLO 凭借其强大的特征提取能力逐渐取代传统方法被应用到遥感图像舰船目标检测中<sup>[24-26]</sup>。由于遥感图像与自然图像存在巨大的差异,经典目标检测算法直接用于舰船检测存在检测精度低、漏检等问题,需要对算法进行改进。

#### 3.1 多尺度舰船目标检测

基于 CNN 的目标检测算法在自然图像数据集上取得了良好的检测效果,而遥感图像中往往背景复杂、舰船目标尺度变化大,经典目标检测算法往往不能有效地提取舰船特征。目前的研究主要采用频域增强、特征金字塔网络结构、注意力机制等方法对目标特征进行增强,进而提高多尺度舰船目标的检测精度。常见方法如图 12 所示。

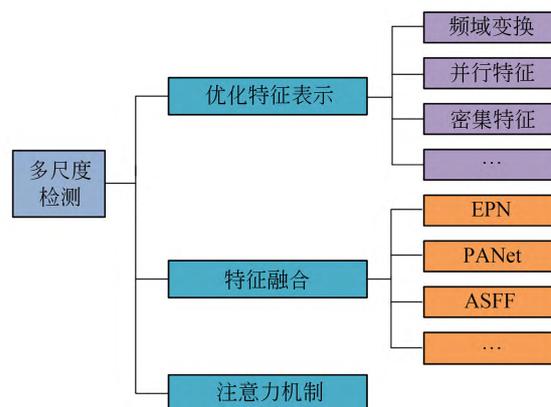


图 12 常用的多尺度检测方法

Fig. 12 Common multi-scale detection methods

##### 3.1.1 优化特征表示能力

###### (1) 频域增强

Al-Saad 等<sup>[27]</sup>提出了频域增强方法,将小波变换嵌入 Faster R-CNN,在提取 ROI 之前,将原始图像分解成高、低频分量,在频域进行训练和测试,提高了检测精度。这种方法简单易行,但是精度提升不高。

###### (2) 并行特征

并行特征可以提高模型的多尺度学习能力,通常对同一输入应用多个不同核大小或扩张速率的卷积来得到。Li 等<sup>[28]</sup>提出了一种分层选择滤波层(Hierarchical Selective Filtering, HSF)对 Faster R-CNN 进行了改进,HSF 由核大小为  $1 \times 1$ ,  $3 \times 3$ ,  $5 \times 5$  的三个并行卷积层组成,通过层次化的卷积运算来生成多尺度舰船特征,有效检测不同尺寸的近岸和近海船舶。Liu 等<sup>[29]</sup>利用  $1 \times 1$ ,  $3 \times 3$ ,  $5 \times 5$  和  $7 \times 7$  的卷积核组成四支结构,并在每个分支中分别引入膨胀率为 1, 3, 5, 7 的膨胀卷积,增加接收野,最后通过残差将四个支路的输出和输入端口连接起来,生成检测多尺度舰船目标的特征。

###### (3) 密集特征

DenseNet<sup>[30]</sup>采用层间密集连接的方法,每层都接受之前所有层的特征图作为额外输入,并将来自不同层的特征图进行拼接,保持底层特征信息的完整性,促进了特征重用,可以提高遥感图像多尺度目标的检测性能。Jiao 等<sup>[31]</sup>提出了一种基于 Faster-RCNN 框架的密集连接多尺度神经网络,该网络将一个特征图紧密地从上到下连接

到其他特征图,并从连接的特征图生成建议,解决了多尺度、多场景合成孔径雷达(Synthetic Aperture Radar, SAR)的舰船检测问题。Tian等<sup>[32]</sup>设计了一种密集特征提取模块,集成不同分辨率的低层位置信息和高层语义信息,提高特征在网络中的重用效率,将该模块应用在经典检测网络YOLO和Mask-RCNN上,改进网络在可见光图像和SAR图像数据集上的检测精度均有提高。

### 3.1.2 特征高效融合

特征融合是提高多尺度目标检测性能的另一常用方法。一般来说,在神经网络中,浅层特征包含目标更多的结构和几何信息,这有利于目标的回归。高级特征包含更多的语义信息,有利于对象的分类。高效的特征融合方法可以提高网络的表征能力,从而提高模型检测多尺度目标的精度。

FPN<sup>[33]</sup>自顶向下将高层强语义信息的特征传递下来,但是对定位信息没有传递。PANet<sup>[34]</sup>在FPN的基础上添加了一个自底向上的金字塔,将低层的强定位信息特征传递上去,融合的特征增添了语义信息。自适应空间特征融合<sup>[35]</sup>(Adaptive Spatial Feature Fusion, ASFF)将每层信息融合起来并且自主学习各个尺度融合的权值,解决了PANet特征融合时小目标在高层特征层上被当作背景和大目标在底层中被当作背景的问题。空洞空间卷积池化金字塔<sup>[36]</sup>(Atrous Spatial Pyramid Pooling, ASPP)对输入以不同采样率的空洞卷积并行采样,将结果进行堆叠,再通过 $1 \times 1$ 卷积将通道数降低到预期数值,以多个比例捕捉图像的上下文信息。

这些特征融合模块被应用于舰船检测网络中<sup>[32,37-40]</sup>,融合不同层次的特征,在保证位置信息准确性的同时保留更多的语义信息,提高多尺度目标的检测效果。Tian等<sup>[32]</sup>引入FPN和ASPP结合的模块进行特征融合,获得更大范围的深度语义信息,增强对不同尺度目标特征的提取能力。Zhang等<sup>[37]</sup>利用改进的FPN构建SAR舰船检测网络,取得了良好的检测效果。Qing等<sup>[38]</sup>利用改进的FPN和PANet对主干网络输出的特征进行融合,加强舰船特征提取。

### 3.1.3 注意力机制

为了优化提取的特征,注意力机制也被融合

到舰船检测网络中<sup>[39-42]</sup>。注意力机制起源于人类的视觉机制,核心思想是关注关键信息而忽略无关信息,减少时间成本和降低计算复杂度。基于视觉注意力机制的目标检测算法通常通过注意力模型得到显著特征图,增强目标与背景的差异,然后通过分析显著特征图对目标进行检测。

注意力机制理论上可以嵌入到网络的任意位置。Chen等<sup>[39]</sup>将注意力模型集成到检测网络的主干部分中,利用注意力模型获得不同层次的显著特征,再用FPN融合不同层次的显著特征。Zhang等<sup>[40]</sup>设计的LSSD(Lightweight Single Shot Detector)在双向特征融合模块进行特征融合后利用注意力机制进一步优化融合特征,有利于更有效地捕捉关键信息。注意力机制的引入提高了SAR图像复杂场景下多尺度舰船目标的检测效果。Qu等<sup>[41]</sup>利用卷积注意力模块(Convolutional Block Attention Module, CBAM)将辅助网络连接到YOLOv3主干网络,使网络能够更好地学习特定的目标特征,然后用ASFF取代FPN解决特征融合不足的问题,提高模型的稳定性,在可见光遥感图像数据集上取得了优于YOLOv3的检测效果。

### 3.1.4 小结

优化特征提取网络和高效的特征融合方法能够有效提高检测精度,但是会增加计算复杂度,降低检测速度,精度与速度的平衡是实际应用中需要考虑的重要问题。大多数注意力机制的参数是通过标签和预测值的损失反向传播进行更新,没有引入其他监督信息,因此受到的监督有局限,容易产生过拟合的问题。

## 3.2 多角度舰船目标的精确定位

与自然图像不同,遥感图像以俯视视角拍摄的,舰船目标往往以不同方向分布在图像中,经典的目标检测模型的水平边界框的定位方式,难以对多方向舰船目标进行精准、紧凑的定位。

### 3.2.1 水平边界框精确定位

#### (1)两阶段水平框检测改进算法

为了提高水平边界框检测多角度舰船目标的效果,旋转不变层和多角度锚点被用于改进经典的两阶段目标检测算法。Cheng等<sup>[43]</sup>在RCNN中引入旋转不变层,优化目标函数来训练网络,旋转不变性通过强制旋转前后训练样本的

特征相互映射接近来实现,解决了任意方向目标检测不准确的问题。Li等<sup>[44]</sup>在Faster R-CNN模型中增加多角度锚点处理特征,解决了水平锚点对旋转物体敏感的问题,能有效检测任意方向的目标。

### (2)一阶段水平框检测改进算法

高斯模型被应用于经典一阶段目标检测算法中,可以增加定位坐标的准确性。Hong等<sup>[45]</sup>在YOLOv3模型中引入锚框坐标的高斯参数来预测定位,解决坐标信息不可靠的问题,改进模型可以应对不同分辨率下遥感图像中舰船目标方向和尺寸的变化,稳定性较高。

### (3)密集目标漏检问题

使用水平边界框检测存在密集目标漏检的问题,因为多个舰船倾斜密集排布时,不同目标边界框之间的重叠区域会很大,在进行NMS操作后置信度低的边界框会被丢弃,造成目标的漏检。为了解决这个问题,Chen等<sup>[39]</sup>在后处理中引入Soft-NMS,在NMS中与置信度最大的边界框IOU超过阈值的其余边界框的置信度均被置为0,而Soft-NMS<sup>[46]</sup>将其余边界框的置信度进行衰减,衰减程度与IOU值有关,衰减后置信度大于正确检测阈值的仍当作是目标,减少严重重叠情况下舰船目标的漏检。这种方法可以一定程度上缓解密集目标漏检的问题,但是阈值的设定对结果的影响较大。

#### 3.2.2 旋转边界框精确定位

为了对多方向舰船目标进行精准定位,旋转边界框被引入经典目标检测网络中提升检测效果。

### (1)两阶段旋转框检测改进算法

两阶段目标检测算法通常利用RPN生成水平锚框来预测目标位置,改进算法使用RPN生成旋转锚框。如Yang等<sup>[47]</sup>提出了一种基于Faster R-CNN的舰船目标旋转框检测框架,在RPN中使用尺度、比例和角度3个参数生成旋转锚框来预测舰船目标的最小外接矩形。Koo等<sup>[48]</sup>提出的RBox-CNN同样使用RPN生成旋转锚框,并用宽度/高度距离投影来稳定地预测角度。两阶段检测网络的检测速度慢,难以满足实时性的需求。

### (2)一阶段旋转框检测改进算法

一阶段检测算法将图像输入检测网络直接输出目标水平边界框的三维坐标信息,旋转边界框可以通过增加角度信息得到。黎元文<sup>[49]</sup>和陈俊<sup>[50]</sup>使用旋转矩形框改进YOLOv3模型,在预测三维位置信息的基础上增加了角度信息,同时改进了损失函数以及计算旋转框IOU的方法,提高了多角度并排停靠舰船目标检测的准确率,相比于两阶段的检测网络检测速度也得到了提高。

### (3)角度边界性问题

以上改进算法中的旋转边界框均可用5个参数 $(x, y, w, h, \theta)$ 来表示,如图13所示。其中, $(x, y)$ 表示中心点坐标, $(w, h)$ 表示宽度和高度, $\theta$ 表示旋转角度,指长边与 $x$ 轴(水平轴)的夹角,角度为 $[-90^\circ, 90^\circ)$ 。五参数表示法会带来角度周期性的边界问题: $-90^\circ$ 和 $89^\circ$ 两个边界角度的偏差非常小,模型计算边界处的损失值却会突然增加,使网络学习难度提高。

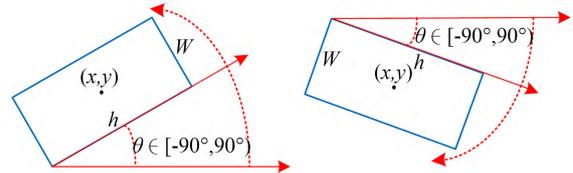


图13 旋转边界框的五参数表示法

Fig. 13 Five-parameter representation of rotating bounding box

为了解决角度边界不连续的问题,Qing等<sup>[37]</sup>将圆形平滑标签(Circular Smooth Label, CSL)技术引入旋转框舰船检测模型中。CSL<sup>[51]</sup>利用高斯函数把连续的目标角度转化为离散类别标签,把回归问题转化为分类问题,如图14所示。Su等<sup>[52]</sup>提出了一种非基于角度的回归方法,取6个参数 $(x, y, w, h, OH, OV)$ 来确定旋转框,如图15所示,其中 $(x, y)$ 表示中心点坐标; $(w, h)$ 表示目标水平外框的宽和高; $H, V$ 分别为水平边界框与顺时针方向旋转边界框之间的水平、垂直距离,然后求出标准化的水平和垂直偏移量: $OH = H/w, OV = V/h$ ,从根本上解决了角度回归的边界性问题。CSL技术和旋转矩形框的六参数表示法都可以有效解决边界问题,但是会增加模型参数量,损失检测时间。

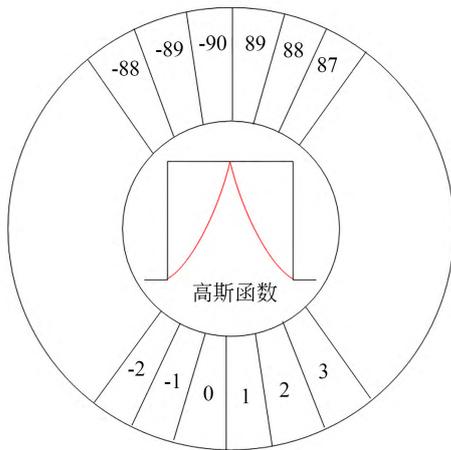


图 14 圆形标签平滑

Fig. 14 Circular label smoothing

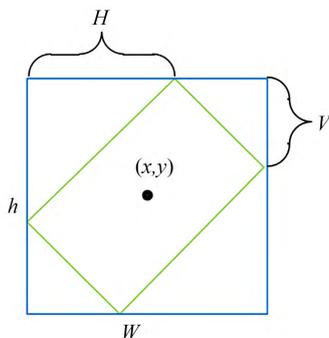


图 15 旋转边界框的六参数表示法

Fig. 15 Six-parameter representation of rotating bounding box

### 3.2.3 无锚框精确定位

无锚框的方法也被应用于旋转舰船目标检测<sup>[53-56]</sup>,如 Wang 等<sup>[54]</sup>提出的基于 CenterNet 的 SAR 图像船舶检测方法,它将舰船目标建模为一个点,回归水平边界框大小,不需要 NMS,从根本上解决了因 NMS 导致的密集目标漏检问题。Cui 等<sup>[55]</sup>提出的一阶段无锚舰船检测框架将旋转舰船目标的检测完全转化为中心关键点和形态大小的预测,并提出了一种“正交池化”模块来提取舰船旋转特征。

### 3.2.4 小结

使用水平边界框检测多角度舰船目标存在密集目标漏检的问题。旋转边界框可以解决漏检问题,并且更好地贴合目标,但是需要考虑角度的边界性问题。无锚框的方法对关键点的预测准确性要求很高,预测不准确就容易导致漏检和定位不准。

### 3.3 提高小目标检测效果

与自然图像中目标占整幅图像的比例很大不同,遥感图像中舰船目标往往只有几十到几百个像素,属于小目标。其检测主要存在两个问题:一是样本不足,很多舰船数据集图像中的小舰船目标并未被标注出来,缺乏大量数据对模型进行训练;二是小目标在图像中所占像素少,随着 CNN 的前向传播,特征图尺度不断减小,顶层特征图中可能不包括小目标的特征,不能帮助小目标定位。小目标检测一直是现有深度学习算法中的难点,目前有大量针对小目标检测开展的算法优化研究,常用方法如图 16 所示。

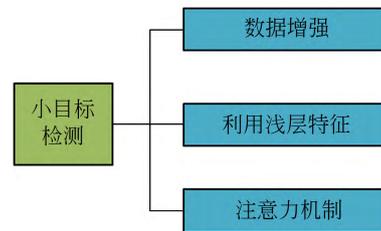


图 16 提高小目标检测效果的常用方法

Fig. 16 Common methods for improving effectiveness of small target detection

#### 3.3.1 数据增强扩充小目标样本

数据增强是提高小目标检测性能的有效方法之一,可以有效地解决小目标样本不足问题。传统的数据增强方法可以大致分为:(1)基于几何变换的方法,包括旋转、缩放、翻转、裁剪、填充、平移和仿射变换等;(2)基于改变颜色成分的方法,包括亮度、对比度、色相和饱和度等。为了增加遥感图像训练数据集的规模和多样性,多种数据增强方法被用于提高检测模型的鲁棒性和泛化能力,常用的技术有多角度旋转、颜色抖动、随机平移、随机裁剪、水平翻转和添加随机噪声等<sup>[57-59]</sup>。

为了解决遥感图像数据集中小型舰船样本稀缺的问题,Shin 等<sup>[60]</sup>提出了一种“剪切和粘贴”策略来增强图像用于训练模型,利用预先训练好的 Mask-RCNN 提取船舶切片,然后粘贴到各种背景海洋场景中,合成新的图像,检测结果验证了合成舰船图像的有效性。Hu 等<sup>[61]</sup>提出了一种混合策略,将海面目标区域与多个变化场景混合,以增加多样性和训练样本的数量。Chen 等<sup>[62]</sup>

提出了一种利用梯度惩罚的高斯混合 Wasserstein GAN 生成足够信息量的小型舰船目标样本,然后用原始数据和生成数据对 CNN 进行训练,实现对小型船舶的精确实时检测。

### 3.3.2 利用浅层特征检测小目标

为了解决小目标在深层特征图中消失的问题,常用方法是充分利用浅层特征中的信息对小目标进行检测<sup>[63-67]</sup>。Kong 等<sup>[63]</sup>提出的 HyperNet 利用跳层提取特征的方式来同时获取包含语义信息的高层特征和包含高分辨率位置信息的浅层特征,利用浅层特征来提高小目标的检测效果。Wei 等<sup>[64]</sup>在 Faster R-CNN 中引入扩张卷积,它可以提供更大的接受域,减少小目标信息的丢失,提高检测效果。Zhang 等<sup>[65]</sup>使用多分辨率卷积改进 Faster-RCNN 的 VGG16 结构,将深层特征和浅层特征映射结合生成多分辨率特征图,提高了小型舰船目标的召回率和准确率。Liu 等<sup>[66]</sup>采用细粒度特征增强对 YOLOv2 模型进行改进,向 YOLOv2 网络中添加重组层和路由层,将前向传播中的浅层特征图和深层特征图汇集起来,提高了小型舰船目标的检测效果。针对 YOLOv3 网络对小型舰船目标检测精度低的问题,常用方法是在网络降 8 倍、降 16 倍和降 32 倍采样 3 个检测尺度的基础上增加 1 个降 4 倍采样的检测尺度,在增加的浅层特征尺度上为小目标分配锚框,提高检测效果<sup>[67-69]</sup>。引入浅层特征会增加模型计算的复杂度,时间成本会增加。

### 3.3.3 注意力机制

注意力机制被引入网络中,通过优化小目标的特征表示来提高检测性能。典型目标检测算法如 YOLOv3, YOLOv4 一般在将多尺度的特征图进行级联后检测目标,从每个通道和位置提取的特征对最终检测结果的贡献平等,但实际上每个通道和位置的特征图是对特定语义信息的响应。注意力机制可以给予代表小目标特征的部分适当的权重,有效提高模型检测小目标的能力。Chen 等<sup>[70]</sup>在 YOLOv3 网络中引入膨胀注意模块(Dilated Attention Module, DAM),它利用膨胀卷积来扩大接收野,并集成通道注意和空间注意模块来提取显著特征,突出小目标与背景的区别,提高检测效果。Nie 等<sup>[71]</sup>在 Mask-RCNN

模型中同时使用通道注意模块和空间注意模块,增强了信息从底层到顶层的传播,提高对小型船舶的检测精度。Hu 等<sup>[72]</sup>为了优化特征信息的表达,提出了空间和通道维度的双注意模块,自适应学习特征在不同尺度上的显著性,并提出了一个新的损失函数,为小目标检测提供了更好的收敛性能。

### 3.3.4 小结

数据增强可以扩充小型舰船目标的样本数量,是提高小目标检测的一项有效措施。利用浅层特征可以提高网络检测小目标的效果,但是会增加计算复杂度、损失检测时间。注意力机制也可以很好地提高小目标检测效果,但是需要考虑过拟合问题。

## 3.4 模型轻量化提高检测速度

典型的目标检测模型参数量巨大,往往难以部署在卫星的嵌入式设备上。为了满足实时检测舰船目标,将模型应用于资源有限的嵌入式设备中,需要减少模型的参数量,提高模型的检测速度,常用方法如图 17 所示。

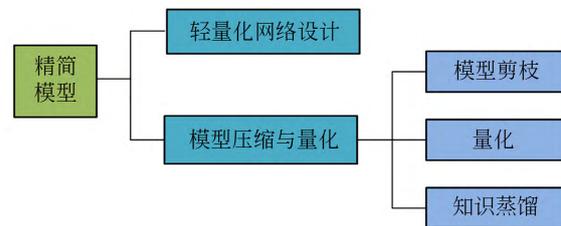


图 17 精简模型的常用方法

Fig. 17 Common methods for streamlining models

### 3.4.1 模型压缩

模型压缩是模型轻量化的一种常用策略。3 种模型压缩方法被广泛应用,分别是模型剪枝、知识蒸馏和量化<sup>[73]</sup>。其中,模型剪枝通过去除网络中冗余的通道或层,可以显著减小参数大小和模型大小<sup>[74-75]</sup>;知识蒸馏将经过训练的大型网络作为教师网络,然后用较大的网络来指导小的学生网络的训练<sup>[76-77]</sup>;量化的核心思想是对网络进行压缩,例如将权值的比特数从 32 位浮点数减少到 16 位浮点数或 8 位整数,使模型尺寸大大降低<sup>[78-79]</sup>。

为了获得更好的舰船目标检测性能,这三种方法通常会结合使用<sup>[80-83]</sup>。Zhang 等<sup>[80]</sup>采用结构

化剪枝方法对网络进行压缩,再采用知识蒸馏来提高压缩后网络的识别精度。Chen 等<sup>[81]</sup>引入基于权重的网络剪枝和权值量化对网络进行压缩。Ma 等<sup>[82]</sup>首先对 YOLOv4 模型进行稀疏训练找到不太重要的信道和层;其次对网络进行信道修剪、层修剪;再利用知识蒸馏对剪枝模型进行再训练;最后,将模型的权值从 FP32 (32-bit Floating Point, 32 位浮点数)量化为 FP16。模型压缩与量化流程如图 18 所示。以上方法对资源受限的 SAR 目标识别都取得了较好的效果,减少了模型参数量,提高了检测速度。陈科峻等<sup>[83]</sup>将 YOLOv3 算法的批量归一化层的尺度因子作为通道重要性的度量指标,对模型进行剪枝压缩,参数量减少了 91.5%,检测时间缩短了 60%,能够满足可见光遥感图像实时舰船检测的需要。

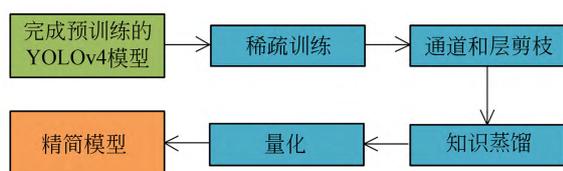


Fig. 18 Model compression and quantification processes

### 3.4.2 网络轻量化

设计轻量化骨干网络结构或卷积计算单元是模型轻量化的另一种有效策略。Li 等<sup>[84]</sup>改进 Faster RCNN 骨干网络结构,提出了一种轻量级舰船探测器 Lite-Faster RCNN,将检测速度提高了 8 倍。Huang 等<sup>[85]</sup>提出的 Ship-YOLOv3 模型通过减少部分卷积运算和增加跳转连接机制改变 YOLOv3 网络结构来减少特征冗余,在保证实时性的前提下模型的检测精度和召回率都得到提高。Ding 等<sup>[86]</sup>提出的舰船检测模型用卷积代替 Faster RCNN 中的全连接层,大大降低了网络参数量,减少了内存需求和时间消耗。Long 等<sup>[87]</sup>结合密集连接、残余连接和群卷积的思想,提出了 Lira-YOLO 舰船目标检测器。与 Tiny-YOLOv3 相比,Lira-YOLO 具有更高的检测精度和更低的计算复杂度。Zhao 等<sup>[88]</sup>在网络中引入深度可分卷积,构建了一个轻量级舰船检测器。

### 3.4.3 小结

模型的压缩与量化虽然可以提高检测速度,但是会损失检测精度,降低模型识别率。网络轻量化可以提高检测速度,降低计算复杂度,不会牺牲检测精度;但过于依赖人工先验知识,需要针对不同场景对网络进行相应的优化,以适应不同的检测任务。

## 3.5 大幅宽遥感图像舰船目标检测

尽管目标检测技术已经取得了长足的进步,但在大幅宽遥感图像中舰船的快速检测仍面临挑战。如果将一幅大幅宽的图像通过降采样直接输入检测网络,图像信息会丢失,不利于检测目标。

### 3.5.1 分块检测

常用的分块检测方法是大幅宽遥感图像分为若干个图像块,分别对每个图像块进行检测识别,检测流程如图 19 所示。Voinov 等<sup>[89]</sup>提出了一种基于卷积神经网络的大幅宽遥感图像舰船检测方法,首先去除陆地区域后将图像分块,然后利用 MobileNet 模型检测图像块中是否包含舰船目标,最后对正分类结果采用 Faster R-CNN 预测舰船的位置和类别。这种图像分块方式容易在图像块的边缘处将目标一分为二,导致目标不完整从而影响检测效果。

基于感兴趣区域提取的图像分块方法可以避免目标被分割的问题。黎经元<sup>[59]</sup>利用形态学方法和视觉显著性算法在大幅宽遥感图像中提取可疑海域切片,避免目标被分割。聂婷<sup>[90]</sup>使用扩展小波变换增强复杂背景下目标与背景的对比度来快速定位大幅宽遥感图像中的感兴趣区域,然后用改进的超复数频域视觉检测方法来提取图像的感兴趣区域。

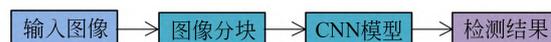


Fig. 19 Large-area remote sensing image segmentation detection process

### 3.5.2 整幅图像一次性检测

Van 等<sup>[91]</sup>提出的 YOLT 将 YOLO 应用于大幅宽遥感图像检测,利用 YOLO 在大幅宽图像上滑动窗口快速检测各个区域,完成对整幅遥感图

像的检测任务。YOLT 滑动窗口时相邻两窗口之间会有 15% 的重叠,最后合并每个窗口检测结果进行非极大值抑制得到最终结果,这种基于滑动窗口的方法重复地处理会占用大量的计算时间,使得整个算法的效率降低,一次性对大幅宽遥感图像进行舰船可以提高检测速度。Su 等<sup>[52]</sup>提出了一种基于 YOLO 的一次性检测大幅宽遥感图像的方法,将特征提取网络改进为全卷积结构组成的 DCNDarknet25,可以接受任何大小的图像作为输入,并且通过减少参数和添加变形卷积提高了检测速度和准确性。

### 3.5.3 小结

大幅宽遥感图像分块检测法第一阶段的检测结果对最终分类结果起决定性作用,利用人工设计的特征提取感兴趣区域,效率低且检测效果不好。一次性检测方法将整幅遥感图像送到网络训练时,大面积的背景信息被当作负样本,小部分的舰船目标被视为正样本,这会导致严重的样本不平衡,浪费训练时间和资源,因此需要一定的策略来解决正负样本不平衡的问题。

## 4 图像数据集和算法性能评价

### 4.1 图像数据集

深度学习需要大量的样本进行模型训练,随着卷积神经网络模型在遥感领域的广泛应用,出现了大规模的遥感图像数据集。包含舰船目标的常见数据集如表 2 所示。其中,最常用于舰船检测的可见光图像数据集是 DOTA 和 HRSC2016, SAR 图像数据集是 SSDD。有研究者收集来源于谷歌地球的遥感图像,构建舰船数据集来训练模型<sup>[28,47,49,66]</sup>。

#### 4.1.1 DOTA 数据集

DOTA 数据集<sup>[92]</sup>由来自不同传感器和平台的 2 806 幅航空遥感图像组成,包括舰船、飞机和桥梁等 15 个类别,其中舰船图像 573 张。图像尺寸在 800×800 到 4 000×4 000 像素之间。数据集总共有 188 282 个实例,舰船目标采用水平边界框与旋转边界框标注。

#### 4.1.2 HRSC2016 数据集

HRSC2016 数据集<sup>[93]</sup>由来自 6 个不同港口的 1 061 幅遥感图像组成。目标包括海上的船舶和停靠在岸边的船舶。图像分为训练集 436 张(共

1 207 个标记样例)、测试集 444 张(共 1 228 个标记样例)和验证集 181 张(共 541 个标记样例)。图像尺寸为 300×300 到 1 500×900 像素,分辨率为 0.4~2 m。舰船目标采用旋转边界框标注。

#### 4.1.3 NWPU VHR-10 数据集

NWPU VHR-10 数据集<sup>[94]</sup>包括一个正面图像集(包括 650 张图像)和一个负面图像集(包括 150 张图像)。NWPU VHR-10 数据集的正像集图像尺寸从 533×597 到 1 728×1 028 像素,包含 10 类地理空间对象:飞机、船舶、储罐、棒球场、网球场、篮球场、田径场、港口、桥梁和车辆。

#### 4.1.4 Kaggle Airbus Ship Detection 数据集

Kaggle Airbus Ship Detection 数据集<sup>[95]</sup>由 Kaggle 挑战赛(Airbus Ship Detection Challenge)提供的 RGB 图像(每幅尺寸为 768×768,包括表示船只位置的编码像素)数据集。编码后的像素被转换为二进制掩码,其中“ship”表示为 1,“no ship”表示为 0。通过计算值为 1 掩码的 4 个角坐标可将其转换为边界框。

#### 4.1.5 MASATI 数据集

MASATI 数据集<sup>[96]</sup>包含 6 212 幅可见光卫星遥感图像,其中舰船图像 3 113 张。图像尺寸约为 512×512 像素。

#### 4.1.6 HRRSD 数据集

HRRSD<sup>[97]</sup>数据集中的图像主要来自谷歌地图,分辨率为 0.15~1.2 m,少部分来自百度地图,分辨率为 0.6~1.2 m。图像共计 21 761 张,包括飞机、棒球场、篮球场、桥梁、十字路口、田径场、港口、停车场、船舶、储罐、T 形路口、网球场和车辆 13 类目标,其中包含舰船目标的图像有 3 886 张,采用水平边界框标注。

#### 4.1.7 DIOR 数据集

DIOR 数据集<sup>[98]</sup>包含 23 463 张图像和 192 472 个实例,涵盖了舰船、飞机、桥梁和篮球场等 20 类对象。其中舰船类图像 2 702 张,分为训练集 650 张、验证集 652 张、测试集 1400 张。图像尺寸为 800×800,分辨率为 0.5~30 m,舰船目标采用水平边界框标注。

#### 4.1.8 FGSC-23 数据集

FGSC-23<sup>[99]</sup>是一个高分辨率光学遥感图像舰船目标精细识别数据集,共有 4 052 个舰船切片,分辨率在 0.4~2 m,包含 23 类舰船目标。

对每个切片标注了类别、长宽比以及舰船方向,按 1:4 将各类图像随机划分为测试集和训练集。

#### 4.1.9 SSDD 数据集

SSDD 数据集<sup>[100]</sup>共有 1 160 张图像,只对像素数大于 3 的舰船目标进行标注,包含 2 456 艘船舶。平均每张图片显示的船只数量为 2.12 艘;按 7:2:1 分为训练集、验证集和测试集 3 部分。图像尺寸为 500×500,分辨率为 1~15 m,舰船目标采用水平边界框标注。

#### 4.1.10 SAR-ship-Dataset 数据集

SAR-ship-Dataset 数据集<sup>[101]</sup>由 102 景高分三号影像 108 景哨兵一号影像组成,包括 43 819 个 256×256 像素的舰船切片,采用水平边界框标

注,同时包括距离和方位。图像分辨率分别为 3, 5, 8 和 10 m。

#### 4.1.11 AIR-SARShip1.0 数据集

AIR-SARShip1.0 数据集<sup>[102]</sup>来源于高分三号卫星拍摄的 31 景图像,按照 2:1 来划分训练集与测试集。图像尺寸为 3 000×3 000,分辨率为 1 m 和 3 m,舰船目标采用水平边界框标注。

#### 4.1.12 HRSID 数据集

HRSID 数据集<sup>[103]</sup>共有 5 604 张图像,由 136 景 SAR 影像裁剪得到,包括小尺寸目标 9 242 个、中等尺寸目标 7 388 个、大尺寸目标 321 个。65% 的图像划分为训练集,35% 的图像划分为测试集,图像尺寸为 800×800,图像分辨率为 0.5~3 m,采用水平边界框标注。

表 2 舰船数据集对比

Tab. 2 Comparison of ship datasets

数据集	图像类型	舰船图像数目	图像尺寸	分辨率/m	标注方式	采用数据集文献
DOTA	可见光	573	800×800~ 4 000×4 000	—	水平边界框、旋转边界框	[24][38][41][48] [50][55][59][68][92]
HSRC2016	可见光	1 070	300×300~ 1 500×900	0.4~2	旋转边界框	[26][39][42][48] [52][55][58][64][93]
NWPU VHR-10	可见光	—	533×597~ 1 728×1 028	0.5~2	水平边界框	[43][44][50][94]
Kaggle	可见光	—	768×768	—	二进制掩码	[27][45][56][71][83]
MASATI	可见光	3 113	512×512	—	水平边界框	[72][96]
HRRSD	可见光	3 886	—	0.15~1.2	水平边界框	[55][97]
DIOR	可见光	2 702	800×800	0.5~30	水平边界框	[42][98]
FGSC-23	可见光	4 052	—	0.4~2	长宽比及方向	[99] [29][31][32]
SSDD	SAR	1 160	500×500	1~15	水平边界框	[37][39][40][50] [81][84][87][100]
SAR-ship-Dataset	SAR	43 819	256×256	3,5,8,10	水平边界框	[37][54][101]
AIR-SARShip	SAR	31	3 000×3 000	1,3	水平边界框	[69][102]
HRSID	SAR	5 604	800×800	0.5~3	水平边界框	[37][53][103]

## 4.2 评价指标

舰船目标检测的常用评价指标有:交并比 (Intersection Over Union, IOU)、精度  $P$  (Precision)、召回率  $R$  (Recall)、平均精度 AP 和均值平均精度 mAP。

### 4.2.1 IOU

IOU 是两个矩形交集面积与两个矩形并集面积之比,如图 20 所示。假设 A 是模型预测框, B 是目标真实框,则:

$$IOU = \frac{area_{A \cap B}}{area_{A \cup B}} \quad (1)$$

一般情况下通过设定 IOU 阈值来判断预测框是否检测到目标物体,即:

$$val = \begin{cases} \text{true}, IOU > T \\ \text{false}, IOU < T \end{cases} \quad (2)$$

其中阈值  $T$  一般取 0.5。

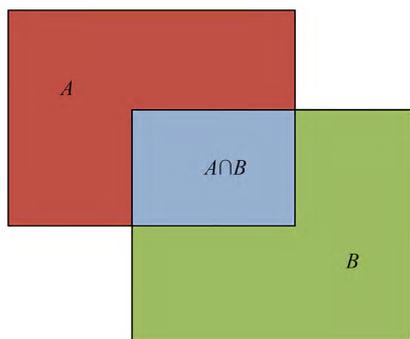


图 20 交并比示意图

Fig. 20 Schematic diagram of IOU

#### 4.2.2 精度和召回率

$P$  表示被正确识别到的正样本数占所有预测为正样本的比例,  $R$  指预测值中被正确识别到的正样本数占所有正样本数量的比例,即:

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$R = \frac{TP}{TP + FN} \quad (4)$$

其中: TP(True Positives)表示预测为正,实际为正; FP(False Positives)表示预测为正,实际为负; FN(False Negatives)表示预测为负,实际为正。

#### 4.2.3 平均精度

PR 曲线是以精度  $P$  为纵坐标、召回率  $R$  为横坐标绘制的曲线,如图 21 所示。模型的精度越高,召回率越高,模型性能就越好,PR 曲线下的面积就越大。

将 PR 曲线下的面积定义为 AP:

$$AP = \int_0^1 P(R) dR \quad (5)$$

AP 的值越大,说明模型的平均准确率越高。

#### 4.2.4 均值平均精度 mAP

mAP 指的是不同类别目标检测精度的均值。在检测多类目标时,计算每一个类别的 AP,然后再计算平均值, mAP 是对检测到的目标平

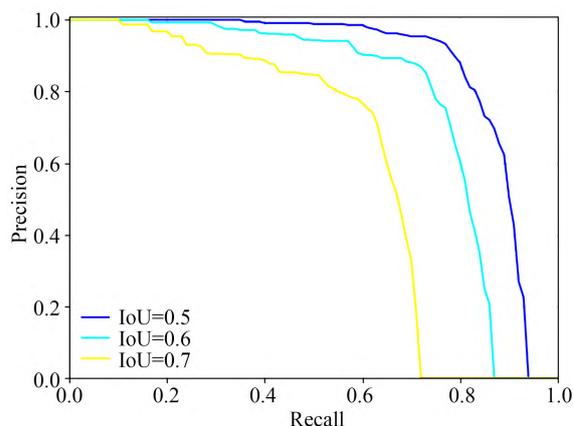


图 21 PR 曲线

Fig. 21 PR curves

均精度的一个综合度量。计算公式如下:

$$mAP = \frac{\sum AP}{m} \quad (6)$$

其中  $m$  表示数据集中目标的类别数目。

## 5 存在的问题和发展趋势

基于深度学习的目标检测技术以其高精度、高效率、适用性强的优良特性广泛应用到遥感图像舰船检测中并取得了一定的效果。但是在实际应用中仍然存在着挑战,主要体现在:

(1) 高分辨率遥感图像中舰船目标清晰可见的同时也会存在大量云雾遮挡等情况,这也是可见光卫星影像的固有缺陷。当图像中有薄云雾以及部分遮挡情况时,可以进行去云雾处理。当图像中有厚云雾遮挡无法检测出目标时,可以融合不同类型传感器的图像信息来抵抗云雾干扰从而完成目标检测任务,如利用 SAR,以及可见光遥感数据进行舰船检测<sup>[104]</sup>。基于多源图像融合的舰船目标检测具有广阔的应用前景,目前已有 SAR 与多光谱/全色图像融合、全色和多/高光谱图像融合、多光谱/高光谱与激光雷达图像融合的研究;多源图像融合需要考虑两方面的问题:一是需要克服不同成像机理差异带来的影响,如 SAR 图像与多光谱/全色图像融合时的噪声和光谱失真;二是需要考虑融合不同来源的地理信息,并从像素级融合往深层次特征级、语义信息融合的方向发展<sup>[105]</sup>。

(2) 遥感图像一般尺寸巨大,现有检测方法

大多采用图像分块的方式,计算复杂,一次性检测方法大面积的背景被当作负样本,正样本数量远远小于负样本数量,训练过程中会出现很多假正例FP,影响检测器性能,正负样本不平衡的问题仍需进一步地研究。在自然图像目标检测中,难分样本挖掘(Hard Example Mining, HEM)是解决正负样本不平衡问题的关键技术。HEM把得分较高的FP当作困难负样本(Hard Negative, HN),并把挖掘出的HN送入网络再训练一次来提高检测器判别FP的能力,从而提高检测精度<sup>[106]</sup>,HEM对大面积遥感图像舰船检测具有借鉴意义。

(3)深度学习目标检测模型的训练需要大量样本,与自然图像动辄百万千万个样本的数据集相比,针对舰船目标的高质量数据集较少,少量的样本训练模型容易出现过拟合。一方面,可采用迁移训练的方法首先在大规模数据集上对模型进行预训练,再用较少的数据集训练模型进行调优;另一方面,可以对数据集进行扩充,然而样本标注需要耗费极大的人力和时间成本。弱监督学习可以显著减少标注工作量,训练集只需要标注二进制标签来标注图像中是否包含目标对象<sup>[107]</sup>,但是二进制标签缺乏位置信息,难以对目标进行定位。针对此问题,Yang等<sup>[108]</sup>通过分析类信息与位置信息之间的相互作用,提出了一种弱监督舰船目标检测器Piston-Net,其检测精度达到了有监督学习目标探测器的水平,然而Piston-Net只能检测单类目标,如何扩展到检测多类目标是今后的研究方向。

(4)当前基于深度学习的舰船目标检测大多是检测图像中是否存在舰船目标并给出其位置。未来的舰船目标检测应更加细化,不只是定位出舰船的位置,还要精准识别出整个编队的配置、各型舰船的数目等,这需要对数据集中的舰船目标进行更细致的标注。FGSC-23是现有的光学遥感图像舰船目标精细识别数据集,但是其样本数量仍然较少,仍需构建用于精细识别的舰船数据集。

(5)虽然基于深度学习的目标检测模型具有很高的检测精度,但是其参数量巨大,实际工程应用需要考虑到实时性检测和硬件设备资源有限需要精简模型。现有的方法有两种:压缩

模型和设计轻量化的网络。前者会损失检测精度,后者可以避免精度损失,但是大多基于现有模型进行精简,未来可以设计智能的模型选择策略,例如神经架构搜索(Neural Architecture Search, NAS)可以从给定的候选神经网络结构集合中按照某种策略搜索出最优网络结构<sup>[109]</sup>,今后研究可以考虑应用NAS自动搜索更优的网络结构。

(6)旋转边界框能够很好地贴合舰船,对目标进行更精细的定位,但是与水平边界框相比,其精度会因边界损失值突变问题而降低。当前,自然图像目标检测中针对边界性问题的解决方法一方面将旋转边界框建模为高斯分布,并提出基于IOU的损失代替斜框IOU损失来简化计算,避免直接角度回归;另一方面把角度预测转化为分类,设计光滑标签,避免边界不连续问题。上述方法已经应用到遥感图像舰船检测中,进一步提高旋转边界框的检测精度仍需进行深入研究,基于中心点或关键点的无锚框旋转检测<sup>[110]</sup>是一个有应用前景的方向。

(7)舰船目标检测未来的发展也应当借鉴计算机视觉等领域最新的研究成果,如基于Transformer的目标检测模型DETR(Detection Transformer),DETR首先用一个CNN网络提取特征,然后展平特征图,当作序列输入给Transformer,经处理后并行输出预测结果;DETR免去了关于锚框的处理,同时免去了NMS后处理<sup>[111]</sup>。但是,收敛速度慢和计算量大等缺点会降低DETR应用于舰船检测任务中的性能。已有学者将基于Transformer的模型进行改进应用于舰船检测,如Zhang等<sup>[112]</sup>提出了舰船目标检测器ESDT(Efficient Ship Detection Transformer)。骨干网采用ResNet50提取深度特征;然后,将特征输入使用多尺度自注意力实现的编码器;最后,增强的特征被送到解码器进行舰船检测;引入了特征蒸馏加快收敛速度,从预训练的大型DETR中学习知识。Chen等<sup>[113]</sup>提出了一种基于PET(Perceptually Enhanced Transformer)的无锚SAR图像舰船检测方法,PET抑制散射噪声,增强舰船目标在复杂背景下的显著特征;并提出一种稀疏注意方法快速聚焦全局特征中的重要信息,加快网络收敛速度。未来,基于Trans-

former 的模型的研究重点是提高检测精度与加速网络收敛。

## 6 结 论

光学遥感图像舰船目标检测是目标检测的

一个重要应用场景,在民用和军用方面都具有重要的价值。本文归纳总结了典型的基于深度学习的目标检测算法,分类梳理了针对舰船目标特点的改进研究,分析了各种改进方法的优缺点。最后,探讨了当前遥感图像船舶目标检测算法面临的问题与挑战以及未来的发展趋势。

### 参考文献:

- [1] 徐芳,刘晶红,孙辉,等. 光学遥感图像海面船舶目标检测技术进展[J]. 光学精密工程, 2021, 29(4):916-931.  
XU F, LIU J H, SUN H, *et al.* Research progress on vessel detection using optical remote sensing image[J]. *Opt. Precision Eng.*, 2021, 29(4): 916-931. (in Chinese)
- [2] 王彦情,马雷,田原. 光学遥感图像舰船目标检测与识别综述[J]. 自动化学报, 2011, 37(9): 1029-1039.  
WANG Y Q, MA L, TIAN Y. Overview of ship target detection and recognition in optical remote sensing images[J]. *Acta Automatica Sinica*, 2011, 37(9): 1029-1039. (in Chinese)
- [3] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]. 2005 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. June 20-25, 2005, San Diego, CA, USA. IEEE, 2005: 886-893.
- [5] 范丽丽,赵宏伟,赵浩宇,等. 基于深度卷积神经网络的目标检测研究综述[J]. 光学精密工程, 2020, 28(5): 1152-1164.  
FAN L L, ZHAO H W, ZHAO H Y, *et al.* Survey of target detection based on deep convolutional neural networks[J]. *Opt. Precision Eng.*, 2020, 28(5): 1152-1164. (in Chinese)
- [6] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C]. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. New York: ACM, 2014: 580-587.
- [7] EVERINGHAM M, ESLAMI S M, GOOL L, *et al.* The pascal visual object classes challenge: a retrospective [J]. *International Journal of Computer Vision*, 2015, 111(1): 98-136.
- [8] HE K M, ZHANG X Y, REN S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [9] EVERINGHAM M, GOOL L, WILLIAMS C K, *et al.* The pascal visual object classes (VOC) challenge [J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [10] GIRSHICK R. Fast R-CNN[C]. 2015 *IEEE International Conference on Computer Vision (ICCV)*. December 7-13, 2015, Santiago, Chile. IEEE, 2016: 1440-1448.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. 2014: *arXiv*: 1409.1556. <https://arxiv.org/abs/1409.1556>.
- [12] REN S Q, HE K M, GIRSHICK R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [C]. *Proceedings of the 28th International Conference on Neural Information Processing Systems-Volume 1*. December 7-12, 2015, Montreal, Canada. New York: ACM, 2015: 91-99.
- [13] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779-788.
- [14] LIU W, ANGELOV D, ERHAN D, *et al.* SSD: Single Shot MultiBox Detector [M]. *Computer Vision - ECCV 2016*. Cham: Springer International Publishing, 2016: 21-37.
- [15] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]. 2017 *IEEE Conference on*

- Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 6517-6525.
- [16] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. 2018: *arXiv*: 1804.02767. <https://arxiv.org/abs/1804.02767>.
- [17] LIN T Y, MAIRE M, BELONGIE S, *et al.* *Microsoft COCO: Common Objects in Context*[M]. Computer Vision-ECCV 2014. Cham: Springer International Publishing, 2014: 740-755.
- [18] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. 2020: *arXiv*: 2004.10934. <https://arxiv.org/abs/2004.10934>.
- [19] ULTRALYTICS. YOLOv5 [EB/OL]. (2020-06-03) [2022-12-01]. <https://github.com/ultralytics/yolov5>.
- [20] LAW H, DENG J. *CornerNet: Detecting Objects as Paired Keypoints*[M]. Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018: 765-781.
- [21] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points [EB/OL]. 2019: *arXiv*: 1904.07850. <https://arxiv.org/abs/1904.07850>.
- [22] ZHU C C, HE Y H, SAVVIDES M. Feature selective anchor-free module for single-shot object detection[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 840-849.
- [23] TIAN Z, SHEN C H, CHEN H, *et al.* FCOS: fully convolutional one-stage object detection[C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27-November 2, 2019, Seoul, Korea (South). IEEE, 2020: 9626-9635.
- [24] 杨阳. 基于深度学习的遥感图像舰船目标检测与分类识别[D]. 武汉: 华中科技大学, 2019.  
YANG Y. *Ship Target Detection and Classification in Remote Sensing Images Based on Deep Learning* [D]. Wuhan: Huazhong University of Science and Technology, 2019. (in Chinese)
- [25] NINA W, CONDORI W, MACHACA V, *et al.* *Small Ship Detection on Optical Satellite Imagery with YOLO and YOLT*[M]. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2020: 664-677.
- [26] 王楠. 基于深度学习的舰船检测识别[D]. 哈尔滨: 哈尔滨工业大学, 2019.  
WANG N. *Ship Detection and Recognition Based on Deep Learning*[D]. Harbin: Harbin Institute of Technology, 2019. (in Chinese)
- [27] AL-SAAD M, ABURAED N, PANTHAKKAN A, *et al.* Airbus ship detection from satellite imagery using frequency domain learning [C]. *Proc SPIE 11862, Image and Signal Processing for Remote Sensing XXVII*, 2021, 11862: 279-285.
- [28] LI Q P, MOU L C, LIU Q J, *et al.* HSF-net: multiscale deep feature embedding for ship detection in optical remote sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(12): 7147-7161.
- [29] LIU S, KONG W, CHEN X, *et al.* Multi-scale ship detection algorithm based on a lightweight neural network for spaceborne SAR images [J]. *Remote Sensing*, 2022, 14(5): 1149.
- [30] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely connected convolutional networks [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 2261-2269.
- [31] JIAO J, ZHANG Y, SUN H, *et al.* A densely connected end-to-end neural network for multiscale and multiscene SAR ship detection[J]. *IEEE Access*, 2018, 6: 20881-20892.
- [32] TIAN L, CAO Y, HE B K, *et al.* Image enhancement driven by object characteristics and dense feature reuse network for ship target detection in remote sensing imagery [J]. *Remote Sensing*, 2021, 13(7): 1327.
- [33] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 936-944.
- [34] LIU S, QI L, QIN H F, *et al.* Path aggregation network for instance segmentation [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 8759-8768.
- [35] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection [EB/OL]. 2019: *arXiv*: 1911.09516. <https://arxiv.org/>

- abs/1911.09516.
- [36] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [37] ZHANG T W, ZHANG X L, KE X. Quad-FPN: a novel quad feature pyramid network for SAR ship detection[J]. *Remote Sensing*, 2021, 13(14): 2771.
- [38] QING Y H, LIU W Y, FENG L Y, *et al.* Improved YOLO network for free-angle remote sensing target detection[J]. *Remote Sensing*, 2021, 13(11): 2171.
- [39] CHEN C, HE C, HU C H, *et al.* A deep neural network based on an attention mechanism for SAR ship detection in multiscale and complex scenarios [J]. *IEEE Access*, 2019, 7: 104848-104863.
- [40] ZHANG X H, WANG H P, XU C A, *et al.* A lightweight feature optimizing network for ship detection in SAR image[J]. *IEEE Access*, 2019, 7: 141662-141678.
- [41] QU Z F, ZHU F Z, QI C X. Remote sensing image target detection: improvement of the YOLOv3 model with auxiliary networks [J]. *Remote Sensing*, 2021, 13(19): 3908.
- [42] CHENG G, LANG C B, WU M X, *et al.* Feature enhancement network for object detection in optical remote sensing images[J]. *Journal of Remote Sensing*, 2021, 2021: 9805389.
- [43] CHENG G, ZHOU P C, HAN J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(12): 7405-7415.
- [44] LI K, CHENG G, BU S H, *et al.* Rotation-insensitive and context-augmented object detection in remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(4): 2337-2348.
- [45] HONG Z H, YANG T, TONG X H, *et al.* Multi-scale ship detection from SAR and optical imagery via A more accurate YOLOv3 [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 6083-6101.
- [46] BODLA N, SINGH B, CHELLAPPA R, *et al.* Soft-NMS-improving object detection with one line of code [C]. *2017 IEEE International Conference on Computer Vision (ICCV)*. October 22-29, 2017, Venice, Italy. IEEE, 2017: 5562-5570.
- [47] YANG X, SUN H, FU K, *et al.* Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks [J]. *Remote Sensing*, 2018, 10(1): 132.
- [48] KOO J, SEO J, JEON S, *et al.* RBox-CNN: rotated bounding box based CNN for ship detection in remote sensing image [C]. *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. November 6-9, 2018, Seattle, Washington. New York: ACM, 2018: 420-423.
- [49] 黎经元. 不同分辨率可见光遥感图像的舰船目标检测识别方法研究[D]. 杭州: 浙江大学, 2021. LI J Y. *Research on Detection and Recognition of Ship Targets from Visible Remote Sensing Images with Different Resolutions* [D]. Hangzhou: Zhejiang University, 2021. (in Chinese)
- [50] 陈俊. 基于R-YOLO的多源遥感图像海面目标融合检测算法研究[D]. 武汉: 华中科技大学, 2019. CHEN J. *Research on Fusion Detection Algorithm of Sea Surface Target in Multi-Source Remote Sensing Images based on R-YOLO* [D]. Wuhan: Huazhong University of Science and Technology, 2019. (in Chinese)
- [51] YANG X, YAN J C. *Arbitrary-Oriented Object Detection with Circular Smooth Label* [M]. Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 677-694.
- [52] SU N, HUANG Z B, YAN Y M, *et al.* Detect larger at once: large-area remote-sensing image arbitrary-oriented ship detection [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 1-5.
- [53] SUN Z Z, DAI M C, LENG X G, *et al.* An anchor-free detection method for ship targets in high-resolution SAR images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 7799-7816.
- [54] WANG X Y, CUI Z Y, CAO Z J, *et al.* Dense docked ship detection via spatial group-wise enhance attention in SAR images [C]. *IGARSS 2020-2020 IEEE International Geoscience and Re-*

- remote Sensing Symposium. September 26-October 2, 2020, Waikoloa, HI, USA. IEEE, 2021: 1244-1247.
- [55] CUI Z Y, LENG J X, LIU Y, *et al.* SKNet: detecting rotated ships as keypoints in optical remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(10): 8826-8840.
- [56] CHEN J J, XIE F Y, LU Y Y, *et al.* Finding arbitrary-oriented ships from remote sensing images using corner detection[J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(10): 1712-1716.
- [57] WEN R, LIU, . An enhanced CNN-enabled learning method for promoting ship detection in maritime surveillance system[J]. *Ocean Engineering*, 2021, 235: 109435.
- [58] WANG Z Q, ZHOU Y, WANG F T, *et al.* SDGH-net: ship detection in optical remote sensing images based on Gaussian heatmap regression[J]. *Remote Sensing*, 2021, 13(3): 499.
- [59] YOU Y N, LI Z Z, RAN B H, *et al.* Broad area target search system for ship detection via deep convolutional neural network[J]. *Remote Sensing*, 2019, 11(17): 1965.
- [60] SHIN H C, LEE K I, LEE C E. Data augmentation method of object detection for deep learning in maritime image [C]. 2020 *IEEE International Conference on Big Data and Smart Computing (BigComp)*. February 19-22, 2020, Busan, Korea (South). IEEE, 2020: 463-466.
- [61] HU J Q, HE J Z, JIANG P, *et al.* SOMC: a object-level data augmentation for sea surface object detection[J]. *Journal of Physics: Conference Series*, 2022, 2171(1): 012033.
- [62] ZHIJUN, CHEN. Deep learning for autonomous ship-oriented small ship detection[J]. *Safety Science*, 2020, 130: 104812.
- [63] KONG T, YAO A B, CHEN Y R, *et al.* HyperNet: towards accurate region proposal generation and joint object detection[C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 845-853.
- [64] WEI S H, CHEN H M, ZHU X J, *et al.* Ship detection in remote sensing image based on faster R-CNN with dilated convolution[C]. 2020 *39th Chinese Control Conference (CCC)*. July 27-29, 2020, Shenyang, China. IEEE, 2020: 7148-7153.
- [65] ZHANG S M, WU R Z, XU K Y, *et al.* R-CNN-based ship detection from high resolution remote sensing imagery[J]. *Remote Sensing*, 2019, 11(6): 631.
- [66] LIU W C, MA L, CHEN H. Arbitrary-oriented ship detection framework in optical remote-sensing images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(6): 937-941.
- [67] CAO C Q, WU J, ZENG X D, *et al.* Research on airplane and ship detection of aerial remote sensing images based on convolutional neural network[J]. *Sensors*, 2020, 20(17): 4696.
- [68] 王浩君. 基于深度学习的光学遥感影像海上舰船目标检测研究[D]. 杭州: 杭州师范大学, 2019. WANG H J. *Research on Ship Target Detection in Optical Remote Sensing Images Based on Deep Learning*[D]. Hangzhou: Hangzhou Normal University, 2019. (in Chinese)
- [69] 韩子硕, 王春平, 付强. 基于深层次特征增强网络的 SAR 图像舰船检测[J]. *北京理工大学学报*, 2021, 41(9): 1006-1014. HAN Z SH, WANG CH P, FU Q. Ship detection in SAR images based on deep feature enhancement network[J]. *Transactions of Beijing Institute of Technology*, 2021, 41(9): 1006-1014. (in Chinese)
- [70] CHEN L, SHI W, DENG D. Improved YOLOv3 based on attention mechanism for fast and accurate ship detection in optical remote sensing images[J]. *Remote Sensing*, 2021, 13(4): 660.
- [71] NIE X, DUAN M Y, DING H X, *et al.* Attention mask R-CNN for ship detection and segmentation from remote sensing images[J]. *IEEE Access*, 2020, 8: 9325-9334.
- [72] HU J M, ZHI X Y, SHI T J, *et al.* PAG-YOLO: a portable attention-guided YOLO network for small ship detection[J]. *Remote Sensing*, 2021, 13(16): 3059.
- [73] DENG L, LI G Q, HAN S, *et al.* Model compression and hardware acceleration for neural networks: a comprehensive survey[J]. *Proceedings of the IEEE*, 2020, 108(4): 485-532.
- [74] LIU Z, LI J G, SHEN Z Q, *et al.* Learning efficient convolutional networks through network slimming[C]. 2017 *IEEE International Conference on*

- Computer Vision (ICCV)*. October 22-29, 2017, Venice, Italy. IEEE, 2017: 2755-2763.
- [75] GUO J Y, ZHANG W C, OUYANG W L, *et al.* Model compression using progressive channel pruning [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31(3): 1114-1124.
- [76] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network [EB/OL]. 2015: *arXiv*: 1503.02531. <https://arxiv.org/abs/1503.02531>.
- [77] ROMERO A, BALLAS N, KAHOU S E, *et al.* FitNets: hints for thin deep nets [EB/OL]. 2014: *arXiv*: 1412.6550. <https://arxiv.org/abs/1412.6550>
- [78] WU J X, LENG C, WANG Y H, *et al.* Quantized convolutional neural networks for mobile devices [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 4820-4828.
- [79] XU Y H, DAI W R, QI Y Y, *et al.* Iterative deep neural network quantization with lipschitz constraint [J]. *IEEE Transactions on Multimedia*, 2020, 22(7): 1874-1888.
- [80] ZHANG F, LIU Y B, ZHOU Y S, *et al.* A lossless lightweight CNN design for SAR target recognition [J]. *Remote Sensing Letters*, 2020, 11(5): 485-494.
- [81] CHEN H Y, ZHANG F, TANG B, *et al.* Slim and efficient neural network design for resource-constrained SAR target recognition [J]. *Remote Sensing*, 2018, 10(10): 1618.
- [82] MA X J, JI K F, XIONG B L, *et al.* Light-YOLOv4: an edge-device oriented target detection method for remote sensing images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 10808-10820.
- [83] 陈科峻, 张叶. 基于YOLO-v3模型压缩的卫星图像船只实时检测 [J]. *液晶与显示*, 2020, 35(11): 1168-1176.
- CHEN K J, ZHANG Y. Real-time ship detection in satellite images based on YOLO-v3 model compression [J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(11): 1168-1176. (in Chinese)
- [84] LI Y D, ZHANG S S, WANG W Q. A lightweight faster R-CNN for ship detection in SAR images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 1-5.
- [85] HUANG H, SUN D C, WANG R F, *et al.* Ship target detection based on improved YOLO network [J]. *Mathematical Problems in Engineering*, 2020, 2020: 1-10.
- [86] DING P, ZHANG Y, DENG W J, *et al.* A light and faster regional convolutional neural network for object detection in optical remote sensing images [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 141: 208-218.
- [87] ZHOU L, WEI S Y, CUI Z M, *et al.* Lira-YOLO: a lightweight model for ship detection in radar images [J]. *Journal of Systems Engineering and Electronics*, 2020, 31(5): 950-956.
- [88] ZHAO Y, YIN Y, GUI G. Lightweight deep learning based intelligent edge surveillance techniques [J]. *IEEE Transactions on Cognitive Communications and Networking*, 2020, 6(4): 1146-1154.
- [89] VOINOV S, HEYMANN F, BILL R, *et al.* Multiclass vessel detection from high resolution optical satellite images based on deep neural networks [C]. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. July 28-August 2, 2019, Yokohama, Japan. IEEE, 2019: 166-169.
- [90] 聂婷. 大幅宽光学遥感图像目标检测技术研究 [D]. 长春: 中国科学院长春光学精密机械与物理研究所, 2019.
- NIE T. *Research on Target Detection Technology of Large-Width Optical Remote Sensing Image* [D]. Changchun: Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, 2019. (in Chinese)
- [91] VAN ETTEN A. You only look twice: rapid multi-scale object detection in satellite imagery [EB/OL]. 2018: *arXiv*: 1805.09512. <https://arxiv.org/abs/1805.09512>.
- [92] XIA G S, BAI X, DING J, *et al.* DOTA: a large-scale dataset for object detection in aerial images [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 3974-3983.
- [93] LIU Z K, YUAN L, WENG L B, *et al.* A high

- resolution optical satellite image dataset for ship recognition and some new baselines[C]. *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods. February 24-26, 2017. Porto, Portugal. SCITEPRESS-Science and Technology Publications*, 2017, 2: 324-331.
- [94] CHENG G, HAN J, ZHOU P, *et al.* Multi-class geospatial object detection and geographic image classification based on collection of part detectors [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 98: 119-132.
- [95] AIRBUS, Airbus Ship Detection Challenge. Kaggle [EB/OL]. (2018-07-30) [2022-08-30] <https://www.kaggle.com/c/airbus-ship-detection>.
- [96] GALLEGO A J, PERTUSA A, GIL P. Automatic ship classification from optical aerial images with convolutional neural networks [J]. *Remote Sensing*, 2018, 10(4): 511.
- [97] CHEN K, WU M, LIU J, *et al.* FGSD: a dataset for fine-grained ship detection in high resolution satellite images [EB/OL]. 2020: *arXiv: 2003.06832*. <https://arxiv.org/abs/2003.06832>.
- [98] KE, LI. Object detection in optical remote sensing images: a survey and a new benchmark[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 159: 296-307.
- [99] 姚力波, 张筱晗, 吕亚飞, 等. FGSC-23: 面向深度学习精细识别的高分辨率光学遥感图像舰船目标数据集[J]. *中国图象图形学报*, 2021, 26(10): 2337-2345.
- YAO L B, ZHANG X H, LÜ Y F, *et al.* FGSC-23: a large-scale dataset of high-resolution optical remote sensing image for deep learning-based fine-grained ship recognition[J]. *Journal of Image and Graphics*, 2021, 26(10): 2337-2345. (in Chinese)
- [100] LI J W, QU C W, SHAO J Q. Ship detection in SAR images based on an improved faster R-CNN [C]. *2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA)*. November 13-14, 2017, Beijing, China. IEEE, 2017: 1-6.
- [101] WANG Y Y, WANG C, ZHANG H, *et al.* A SAR dataset of ship detection for deep learning under complex backgrounds [J]. *Remote Sensing*, 2019, 11(7): 765.
- [102] 孙显, 王智睿, 孙元睿, 等. AIR-SARShip-1.0: 高分辨率 SAR 舰船检测数据集[J]. *雷达学报*, 2019, 8(6): 852-862.
- SUN X, WANG ZH R, SUN Y R, *et al.* AIR-SARShip-1.0: high-resolution SAR ship detection dataset[J]. *Journal of Radars*, 2019, 8(6): 852-862. (in Chinese)
- [103] WEI S J, ZENG X F, QU Q Z, *et al.* HRSID: a high-resolution SAR images dataset for ship detection and instance segmentation[J]. *IEEE Access*, 2020, 8: 120234-120254.
- [104] PARK K A, PARK J J, JANG J C, *et al.* Multi-spectral ship detection using optical, hyperspectral, and microwave SAR remote sensing data in coastal regions [J]. *Sustainability*, 2018, 10(11): 4064.
- [105] 李树涛, 李聪妤, 康旭东. 多源遥感图像融合发展现状与未来展望[J]. *遥感学报*, 2021, 25(1): 148-166.
- LI SH T, LI C Y, KANG X D. Development status and future prospects of multi-source remote sensing image fusion[J]. *Journal of Remote Sensing*, 2021, 25(1): 148-166. (in Chinese)
- [106] SHRIVASTAVA A, GUPTA A, GIRSHICK R. Training region-based object detectors with online hard example mining[C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 761-769.
- [107] ZHOU Z H. A brief introduction to weakly supervised learning [J]. *National Science Review*, 2018, 5(1): 44-53.
- [108] YANG Y, PAN Z X, HU Y X, *et al.* PistonNet: object separating from background by attention for weakly supervised ship detection [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 5190-5202.
- [109] ELSKEN T, METZEN J H, HUTTER F. *Neural Architecture Search*[M]. Automated Machine Learning. Cham: Springer International Publishing, 2019: 63-77.
- [110] WANG X, WANG G, DANG Q, *et al.* PP-YOLO-R: an efficient anchor-free rotated object detector [EB/OL]. 2022: *arXiv: 2211.02386*. <https://arxiv.org/abs/2211.02386>.

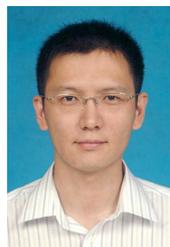
- [111] CARION N, MASSA F, SYNNAEVE G, *et al.* *End-to-end object detection with transformers* [M]. Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 213-229.
- [112] ZHANG Y N, ER M J, GAO W X, *et al.* High performance ship detection via transformer and feature distillation [C]. 2022 *5th International Conference on Intelligent Autonomous Systems (ICoIAS)*. September 23-25, 2022, Dalian, China. IEEE, 2022: 31-36.
- [113] CHEN Y Y, XIA Z H, LIU J, *et al.* TSDet: end-to-end method with transformer for SAR ship detection [C]. 2022 *International Joint Conference on Neural Networks (IJCNN)*. July 18-23, 2022, Padua, Italy. IEEE, 2022: 1-8.

## 作者简介:



黄泽贤(1999—),男,河南濮阳人,硕士研究生,2020年于郑州大学获得学士学位,主要从事遥感影像智能处理、工程仿真等方面的研究。E-mail: huangzexian@mails.ucas.ac.cn

## 通讯作者:



姜肖楠(1981—),男,吉林市人,博士,研究员,博士生导师,2010年于哈尔滨工业大学获得博士学位,主要从事空间光学遥感相机总体技术、遥感影像处理等方面的研究。E-mail: jxn\_cioomp@qq.com